

Residential Location Choice
- The influence of migration and commuting distance -

H.P. Blijie
Delft University of Technology
Faculty of Civil Engineering
Department of Transport and Planning

h.p.blijie@citg.tudelft.nl

Bijdrage aan het Colloquium Vervoersplanologisch Speurwerk 2004,
25 en 26 November 2004, Zeist

Table of contents

1	Introduction.....	1
2	The impact of transport on residential choice: literature review.....	2
3	Conceptual model for residential choice	3
3.1	<i>Residential mobility.....</i>	3
3.2	<i>Conceptual model</i>	4
3.3	<i>Accessibility in the model.....</i>	4
4	Data to use	5
5	Residential utility function with migration distance and commuting time	5
5.1	<i>Observed moves in the WBO2002.....</i>	5
5.2	<i>The distance decay functions for migration distance and commuting distance</i>	7
5.3	<i>Derivation of the utility function.</i>	9
6	Estimation of the simultaneous influence.....	11
6.1	<i>Initial choice set generation: residential search area.....</i>	11
6.2	<i>Estimation of the MNL</i>	13
6.3	<i>Conclusion estimation MNL.....</i>	14
7	Resume & Conclusion.....	14
7.1	<i>Next research steps</i>	15
	Acknowledgements.....	15
	References.....	15

Samenvatting

Woonlocatiekeuze – de invloed van verhuis- en woonwerkafstand.

Een belangrijk onderdeel van geïntegreerde landgebruik-transportmodellen is het modelleren van de deelsystemen zoals het verkeerssysteem, de vastgoedmarkt en het locatiekeuzegedrag van bedrijven en huishoudens. Deze paper gaat over het modelleren van het woonlocatiekeuzegedrag van huishoudens en met name de invloed van bereikbaarheid daarop. In de paper is een eerste aanzet gedaan naar het meenemen van persoonlijke eigenschappen en relaties in een discreet keuzemodel voor woonlocatiekeuze. In eerste instantie is de separate invloed van verhuis- en woonwerkafstand bepaald door het schatten van afstandvervalfuncties voor beide relaties. Hieruit blijkt dat een negatieve machtsfunctie de beste benadering geeft. Uit de statistische verhuisgegevens blijkt echter dat beide relaties simultaan werken bij het bepalen van de woonlocatie. Door middel van een Multinomiaal Logit Model is vervolgens deze simultane werking geschat. De uitkomsten van de modelschattingen laten zien dat een, bij gebruik van een negatieve power functie, de invloed van de migratieafstand bijna twee keer zo groot is als van de woonwerkafstand. Daarnaast resulteert een toename in de verhuisafstand van 1% een afname van bijna 2% dat die locatie gekozen wordt.

Summary

Residential Location Choice - The influence of migration and commuting distance.

An important part of integrated land-use and transport models is the modelling of subsystems like the transport system, the real estate market and the location choice behaviour of households and firms. This paper is about modelling the residential choice behaviour of households and in particular the influence of transport on this. A first attempt is made to incorporate personal attributes and relations in a discrete choice model for residential choice. Firstly the separate influence of migration and commuting distance is determined by estimating three distance decay functions for both relations. It turns out that a negative power function gives the best fit. However, from the statistical migration figures it is shown that migration and commuting distance work simultaneous when deciding where to settle. This simultaneous influence is estimated through a Multinomial Logit model. The results of these estimations show that when using a negative power, the influence of migration distance on residential location choice is almost twice as strong as the commuting distance. Also it turns out that an increase in migration distance of 1% reduces the chance of an alternative to be chosen with almost 2%.

1 Introduction

Many researchers have indicated the interaction between land-use and transport as important knowledge to address policy issues in the field of transport and spatial planning (Miller et al., 1999; Wegener et al., 1999). Recently, more effort has been put in developing more advanced and realistic models that simulate the development of (urban) regions. This regained interest is primarily initiated by three factors, namely: the environmental issues our modern societies have to face, the new methodological insights and the improved computer power of desktop systems. To meet this interest, the Land-Use and Transport interaction models (LUTI-models) have been developed. LUTI-models combine a “traditional” transport with a land-use model. This connection is based on the mutual influence of Land-use and transport (Wegener et al., 1999). Within this line of reasoning it is assumed that the spatial distribution of functions generates activities, which result in flows of people and goods. These flows load the transport network and change travel times and therefore accessibility through congestion. Accessibility is then regarded as an attractiveness of locations, and therefore has an impact on land-use.

The conceptual framework of LUTI-models consists mainly of four different urban markets, namely the land market, the real estate market, the housing market and the labour market. On these markets three groups of so-called spatial agents act, i.e. households, firms and governments (Miller et al., 1999; Wegener et al., 1999).

This paper focuses on one specific part of the urban market, namely the residential choice behaviour of households and its relationship with accessibility. After all, accessibility is often seen as the link between transport and land-use, and in this case between the transport system and the residential choice of households. However, as we will see from the literature overview, this relationship is hard to prove empirically. The paper exists partly of the description of a conceptual framework on how this choice process can be modelled and partly on empirical results of a discrete choice model estimated on empirical migration data. With this model we show the relative influence of migration and commuting distance on residential location choice.

The paper is organised as follows. First we discuss the literature on the impact of accessibility on residential choice (chapter two). Then we present the conceptual model for residential location (chapter three) and the data we plan to use (chapter four). Next in chapter five, distance decay functions are estimated for the separate migration and commuting distance and the derivation of a utility function is given, including these distances. The parameters of these utility functions are estimated by a Multinomial Logit Function in chapter six, using an ellipse formed search area. Finally, chapter seven rounds the paper up with a short summary of the most important conclusions and several future research steps.

2 The impact of transport on residential choice: literature review

This part of the paper gives a brief overview of the conceptual and empirical research on the influence of transport on residential choice that has been conducted in the past decennia. It discusses into several issues varying from the early researches to the latest progresses in residential choice modelling.

The relationship between transport and land use has always been an appreciated subject of study amongst researchers and policymakers who are active on the overlapping area of mobility and space. The earliest researchers of land use (Weber, 1909; Christaller, 1962; Alonso, 1964; Lowry, 1964) incorporated the direct link between land-use and transport into the descriptions of their (visions on) spatial models. Although through the years more nuanced views have risen, it is still assumed that a large, mutual influence exists (see, for example, the circle of Wegener, (Wegener et al., 1999)). Now, the influence of land-use on transport is, or rather: seems, evident. The spatial distribution of people and activity places generates flows between places in space and time. And although predicting the occupation of a transport network through a transport model is not a sinecure, the causal relation is established (Hensher & Button, 2000). The influence of accessibility on land-use and the influence of accessibility on residential choice are less obvious to measure. The following paragraphs give an overview of some of the empirical researches into residential choice.

The early spatial models tended to describe already the models relation between residential choice and transport was quite evident. In this the residential location choice of households was determined by the price of the dwelling and the distance to the work location, concentrated in de Central Business District. The classical residential location choice models of Lowry (1964), Alonso (1964), Muth (1969) describe these phenomena.

As the cities' complexity grew, new ways of describing the residential location choice behaviour arise. Most methodological techniques stem from the field of econometrics. Clark & Van Lierop (1986) have given an overview of the state of the art in modelling residential choice at that time. However, the then available armamentarium is still the basis of most of the (empirical) studies. The two most applied techniques are the Random Utility Theory (RUT) and the Hedonic Price Method.

One of the first attempts to apply the RUT on residential choice is described in McFadden (1978). Despite the fact that this research lacks empirical results, it is one of the first publications on the application of the then still fairly new discrete choice technique on residential choice. The role of "general" accessibility, like the mass of labour, people or services within 30 minutes, seems difficult to quantify in residential choice models. (Molin & Timmermans, 2003). Waddell (1996) even finds a negative, or in other cases insignificant relation between residential location choice and accessibility of jobs and inhabitants. This is probably due to the fact that in this case these measures give more an indication of the preference for the amount of urbanism or urban density. Srour et al. (2002) observe a positive influence of an (logsum) accessibility of jobs on residential location choice. This study, then again, incorporates accessibility measures only, so nothing can be said about substitution effects with other residential choice factors like the characteristics of the dwelling and the neighbourhood. These still prove to be more significant and influential than accessibility.

The relative weak relation between general accessibility and residential choice could be caused by two issues, namely: an insufficient amount of variation between regions and the correlation between variables in revealed preference data. Especially in highly urbanised areas the difference in accessibility between locations is little. The other cause for not successive outcomes of residential models is that they mainly use Revealed Preference Data (RPD),

which generates a lot of correlation between (explaining) variables. After all, many variables in RPD describe more or less the same phenomenon, e.g. the price per square meter and accessibility of people as a degree of urbanisation.

On the other hand, commute distance is observed several times as an influential factor (Evers, 1990; Weisbrod et al., 1980; Moilin & Timmermans, 2003), as well in models that study the behaviour of two-earner households (Timmermans et al., 1992; Rich & Nielsen, 2001, Waddel, 1996). Recently, new insights look more at a household's unique situation, by taking along personal, spatial relations (e.g. work location, position of the kids' school) and how these relate to daily activity patterns and long-term location choices (Axhausen et al. 2001; Olatubara, 1998).

A general introduction and first application of the Hedonic Price Method to determine the price of a dwelling or a residential location can be found in Rosen (1974). Since then, many follow-ups have been made. Recently, Pagliara & Preston (2003) and Srour et al. (2002) investigated the relation between housing costs and accessibility, and came up with a positive relation.

Other methods to study the impact of accessibility on residential choice are scarce. A psychometric analysis by Benjamin & Paaswell (1981) showed that transport and accessibility play only a limited role. The elimination-by-aspects method, used by Young (1984), demonstrated a strong preference for accessibility of schools and shops, a threshold value for commuting and a limited influence of congestion and nearness of friends.

In conclusion we can say that, although a respectable amount of empirical studies have been conducted, the outcomes are rather disappointing. The most promising directions concern accessibility measures that take into account the personal situation of the household, like the daily activity pattern and the therewith linked spatial network of a household.

3 Conceptual model for residential choice

In Blijie (2004) there is a more elaborate description of the conceptual model that we are using this research, including some references to support it. Here we will only discuss the most important issues, namely the importance of residential mobility for this research, the conceptual model summarised and the way accessibility is taken along.

3.1 Residential mobility

The two driving forces behind move decisions are the life cycle of households and changes in the behaviour on the labour market of one or more members of the household (Rossi, 1955). The first motivation leads usually to short distance migrations, the second (labour) to longer. Rossi (1955) and Clark & Dieleman (1996) discuss this relationship between the life cycle of a household and its residential behaviour. The life cycle of a household consists roughly of four steps: change/origin household – dissatisfaction current dwelling – search for a new dwelling – move. The two most important steps of the life cycle for our research are the search and the move to a new dwelling.

When searching for a new dwelling, one can assume that given the constraints of the household and the available supply of dwellings, a household strives for a certain type of utility maximisation or, at least, satisfaction (Rouwendal, 1990). According to the concept of utility maximisation, a household seeks for the dwelling which attributes all meet the preferences in the best way. When this is not possible, simply because such house doesn't exist, substitution behaviour takes place (Hooimeijer, 1990).

A household will only move if it finds a suitable dwelling. Suitable here means above a certain minimal amount of utility or the utility received from the current dwelling. If it cannot find a suitable dwelling, the household does not move. It will then wait until a better occasion arises, change its preferences or adapt the current dwelling in such a way that the direct need to move disappears (Priemus, 1984).

3.2 Conceptual model

The conceptual model for residential search behaviour consists of three steps, and it is designed to take account for the individual processes and characteristics of a household. In this way we hope to identify the unique preferences for the residential situation in general, and for accessibility in particular.

The three steps are:

- a. Specific household groups. In this step we try to categorize a household into a certain group of households, which we assume to have common, specific behaviour regarding residential, mobility or lifestyle preferences.
- b. Systematic choice set generation. When a household is categorised, a feasible choice set is created based on the attributes of the household itself and the group. This set can be determined by spatial, financial, physical or other restrictions.
- c. Disaggregated choice model. Finally a discrete choice model is used to model the household's choice for a dwelling and its location as a 4-digit postal zone.

3.3 Accessibility in the model

In all the mentioned steps of the model, accessibility is incorporated, sometimes more directly than other. To understand the influence of accessibility on the residential choice process, we will try to distillate the importance of accessibility. This means that we try to filter out the impact of other location factors.

In this context it is important to acknowledge the difference between generic and individual accessibility. The first is equal for every (type of) person on a certain location, e.g. the number of jobs available, the distance to the nearest train station, etc. This type of accessibility measures, as can be seen from the literature, turns out to be weak explaining variables.

When we look at individual measures, we notice a strong relationship with the personal network of daily and weekly activities. This is best represented by the influence of the migration distance and commuting time on the migrations, as we will see in the next section of this paper. Given that the budget of a household for living is usually set (and optimally used), this household can make the decision whether to choose for a smaller apartment downtown or a larger semidetached house in the suburbs, i.e. a more versus a less accessible location. In terms of individual accessibility, as the concentration of jobs is still predominantly in the CBD's, the majority of the households will decide either to commute and live in a larger house, or to live small closer to the workplace.

Of course, the personal social network brings along other activities as well, that can be represented by more general accessibility measures. The importance of these accessibilities, however, is assumed to differ per household type. For example, a younger couple without children are more attracted to a downtown area, where many commercial services are situated, whilst a family with younger kids are looking for a place with a primary school and children day-care centre nearby.

In the next sections, we will present a first attempt to construct a discrete choice model by incorporating just the migration and commuting. But first we will discuss the data we are using in this research.

4 Data to use

The most important data source is the Housing Demand Survey (in Dutch: Het Woningbehoefte Onderzoek; WBO). This is a large survey that is kept every two years under a respectable amount of households. The last edition (WBO2002) consists of over 100.000 interviews, representing an equal sample of the Dutch population. In these interviews people were asked about their future housing wishes and former moving behaviour along with questions on household characteristics like age, income, education, labour situation, etc. Also other issues come across like mobility behaviour and leisure activities.

The supply of vacant dwellings from which the households choose, is only available at an annual base from the SYSWOV database (Housing Supply System, in Dutch: Systeem Woningvoorraadgegevens). From this record a quarterly supply can be generated. How this is done will be presented in future papers.

The travel behaviour is monitored by the Dutch National Travel Survey (in Dutch: Onderzoek Verplaatsingsgedrag; OVG). For this survey 140.000 persons from 60.000 households were asked to fill in a diary on their trips for a week. This resource could supply the information on travel preference that is needed for this research. Finally, the factor and cluster analyses to distillate the destination groups are carried out on the WBO and/or OVG analysis.

5 Residential utility function with migration distance and commuting time

In this section we would like to make a first attempt to include personal accessibility measures by exploring the influence of the migration distance and commuting time of the head of the household. To do so, we will first discuss brief the migration data and the migration distance and commuting time. Then we derive two distance decay functions. Eventually, we compose a utility function where both functions are combined.

5.1 Observed moves in the WBO2002

As mentioned when we discussed the data, in the National Housing Survey (WBO2002) over 100.000 people were questioned. From these around 11.500 people were moved in the past two years (2000-2002). Finally, almost 8.000 of these movers also had a working relation. The level of detail of the information on the current and previous residential location is a four digit postal zone. In the Netherlands, the average size of a postal zone is nine km². This size varies from a few square kilometres in urban areas to 100 km² in rural regions. For the work location only the name of the town is known. The sizes of these towns vary from large cities (e.g. Amsterdam) to small villages.

To analyse the migration distance and the former and current commuting time of movers, we added a travel time to the moving and commuting relations. For the commuting travel time we used congested, morning peak travel times derived from the National Model System (in Dutch: Landelijk Model System, LMS), a renowned transport model in the Netherlands. The migration distance is calculated with off-peak travel times, representing the

time for social activities. Table 1 gives some statistical figures of the migration distance, the current commuting time and the change in commuting time due to the move.

TABLE 1: Statistical figures migrations WBO2002.

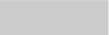
Distance (minutes)	Migration	Commuting	Change commuting
Mean	18,9	26.5	- 5.9
Standard deviation	28,4	26.2	- 27.9
Percentiles			
50	8.5	18.1	0,0
85	33.1	49.0	- 12.0
95	83.6	72.8	- 60.8
Number observations	11652	7802	7783

From these numbers it can be seen that a substantial part of the moves take place on a fairly short distance: over half of them are within 10 minutes, and 85 percent under 33 minutes. As for the commuting time, people are more flexible. The relation between the median and the standard deviation also shows the skewed distribution of the migration distance and commuting time, something that is known from the literature.

When we look at the cross table 2, we see that, as expected, households in general have an aversion against having both migration distance and commuting time. The dislike against moving however is stronger. That is, people tend to move to a location close to the old dwelling, while keeping the commuting distance within a certain distance.

TABLE 2: Percentage share moves to migration distance and (new) commuting time.

		Migration distance (min.)							Total
		0 – 5	5 – 10	10 – 20	20 – 40	40 – 80	80 – 160	> 160	
Commuting time (min.)	0 – 10	12.5%	7.5%	4.4%	1.8%	1.4%	1.0%	0.2%	28.8%
	10 – 20	8.3%	5.2%	6.6%	2.2%	1.6%	1.0%	0.1%	24.9%
	20 – 30	4.8%	2.5%	3.5%	2.0%	0.9%	0.8%	0.0%	14.5%
	30 – 45	4.3%	2.9%	2.9%	2.8%	1.2%	0.5%	0.0%	14.7%
	45 – 60	2.2%	2.0%	2.0%	1.2%	1.1%	0.3%	0.0%	8.8%
	60 – 90	1.3%	1.0%	0.9%	0.7%	1.0%	0.3%	0.0%	5.2%
	> 90	1.2%	0.6%	0.3%	0.3%	0.3%	0.5%	0.0%	3.2%
	Total	34.6%	21.6%	20.7%	10.9%	7.5%	4.3%	0.4%	100%

 : sum = 74.2%

On the other hand, table 3 shows that, when a household migrates over a longer distance, it decreases its (large) commuting distance. This implies that it leaves the old residential location to settle somewhere else near the job place.

TABLE 3: Percentage share moves to migration distance and decrease in commuting time.

	Migration distance (min.)						
	0 – 5	5 – 10	10 – 20	20 – 40	40 – 80	80 - 160	> 160
> 90							
60 – 90					0.3%	0.2%	0.0%
45 – 60				0.0%	0.5%	0.0%	
30 – 45			0.1%	0.8%	0.7%	0.0%	
20 – 30			0.3%	1.2%	0.2%	0.0%	
10 – 20	0.0%	0.1%	2.2%	1.5%	0.2%	0.0%	
0 – 10	3.3%	10.9%	7.8%	1.3%	0.2%	0.1%	
0	28.2%	0.1%	0.0%		0.0%		
-10 – 0	3.2%	10.2%	8.0%	1.4%	0.3%	0.0%	
-20 – -10		0.0%	2.2%	2.0%	0.4%	0.1%	
-30 – -20			0.2%	1.8%	0.5%	0.1%	
-45 – -30			0.0%	0.9%	1.1%	0.1%	
-60 – -45				0.1%	1.8%	0.1%	0.0%
-90 – -60					1.3%	1.2%	0.0%
< -90					0.1%	2.4%	0.4%

█ : sum = 76.0%

5.2 The distance decay functions for migration distance and commuting distance

Tables 2 and 3 show that the migration distance and the commuting time both have an influence on the residential location choice. Analogous to transport studies, this influence can be described as a distance decay function where the chance of a trip is subject to the distance of the trip. In this context: the chance that a household settles on a location is subject to the migration distance and the (new) commuting time. This chance is calculated as the number of (observed) moves per time class (of five minutes), divided by the number of possible moves in that class. In fact, the distance decay function is derived from the spatial interaction function, which describes the amount of trips (or moves) made through the amount of possible trips. In equation 1, the distance decay function $F(d_j)$ is given.

$$F(d_{ij}) = \frac{T_{ij}}{X_i \cdot Y_j} \quad (1)$$

in which: T_{ij} = the interaction between i and j
 X_i = the number of activities in i
 Y_j = the population in j

In an overview of distance decay functions, given by Geurs and Ritsema van Eck (2001), a distance decay function can be described by several functions, namely: a negative power, a negative exponential, a log-logistic and a modified normal (Gaussian). Considering the shapes of the measured distance decay functions, we decided to investigate which of the first three functions fits the data best. The formulations are as follows:

$$\text{- negative power: } F(d_{ij}) = \beta \cdot d_{ij}^{-\alpha} \quad (2)$$

$$\text{- negative exponential: } F(d_{ij}) = \beta \cdot e^{-\alpha \cdot d_{ij}} \quad (3)$$

$$\text{- Log-logistic: } F(d_{ij}) = \frac{\beta}{1 + e^{\alpha + \gamma \cdot \ln d_{ij}}} \quad (4)$$

with α , β and γ as the function parameters and d_{ij} as the distance between point i and j .

The estimations of each formulation for the estimated data are shown in figures 1 and 2.

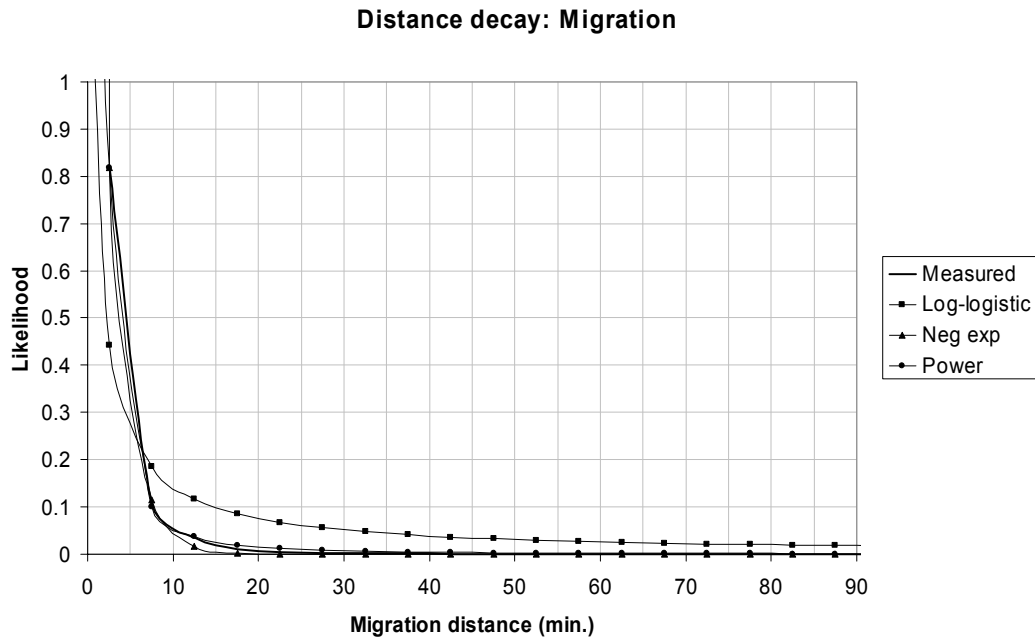


FIGURE 1: Distance decay functions for migration.

Distance decay: Commuting

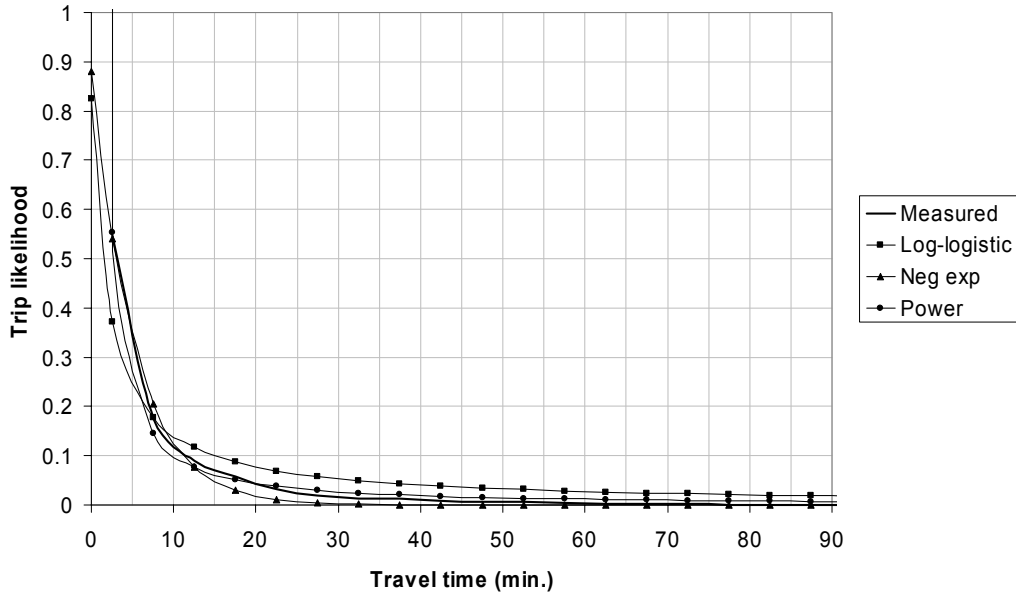


FIGURE 2: Distance decay functions for commuting.

From the function estimators, it appears that the estimated power function approximates the measured migration and commuting distances best. The estimations of the power function gave the following values for the parameters α and β : respectively 1.912 and 4.724 for migration, and 1.695 and 1.221 for commuting. Fotheringham and O’Kelly (1989) support the usage of a power function for long-distance migrations, but our results seem to prove that it also fits best for shorter distances.

However, these power-functions (with these parameters) cannot be used to calculate the chance that a residential location is chosen. After all, the location choice is a simultaneous choice in which both relations play a part and are non-independent. That is to say, the two estimated functions should not be “simply” added up (or multiplied) to retrieve the chance; they have to be estimated simultaneously. To do so, we have chosen to compose a utility function that combines both distance decay functions, as can be seen in the next section.

5.3 Derivation of the utility function.

In this section we will specify a utility function that can be used to estimate the simultaneous influence of the migration distance and commuting time. The distance decay function that showed the best fit on the empirical data (negative power) will be used for this. For further research however, the retrieved function can also be extended with more location- and household specific variables.

The simplest form of a Logit-function, the Multinomial Logit Model (MNL), with the chance $P_c(i)$ that an option i is chosen from a subset c of all alternatives is:

$$P_c(i) = \frac{e^{V_i}}{\sum_{k \in c} e^{V_k}} \quad (5)$$

with for the utility V of option i :

$$V_i = \theta_1 \cdot v_{i1} + \dots + \theta_n \cdot v_{in} \quad (6)$$

with: v = partial utility of an option
 θ = parameters of the partial utility

At first, only the migration distance and commuting time will be included. According to Nijkamp and Reggiani (1992), the logarithm of the distance decay function F represents the (dis-)utility of a trip or move, so we can write:

$$V_i = \theta^o \cdot \ln F^o(d_{oi}) + \theta^w \cdot \ln F^w(d_{iw}) \quad (7)$$

with: d_{oi} = distance between old dwelling o and optional, new dwelling i
 d_{iw} = distance between optional, new dwelling i and workplace w
 θ_o = parameter of the distance d_{oi}
 θ_w = parameter of the distance d_{iw}

Since both distance decay functions F^o and F^w (the likelihood of a move respectively trip) are decreasing with the distances (d_{oi} and d_{iw}), they have a negative influence on the utility of a (new) residential location for larger distances. This means that we expect both θ_{oi} and θ_{iw} to be positive.

When we combine expression (7) with the functions for the distance decay (2), (3) and (4), we get three equations for the negative power, negative exponential and log-logistic form. In these functions the O and W indicate the parameters regarding respectively migration and commuting. The three utility functions are derived below.

For the negative power function, combining (7) and (2) gives:

$$V_i = \theta^o \cdot \ln(\beta^o \cdot d_{oi}^{-\alpha^o}) + \theta^w \cdot \ln(\beta^w \cdot d_{iw}^{-\alpha^w}) \quad (8)$$

which can be rewritten as:

$$V_i = \theta^o \cdot \ln \beta^o - \alpha^o \cdot \theta^o \cdot \ln(d_{oi}) + \theta^w \cdot \ln \beta^w - \alpha^w \cdot \theta^w \cdot \ln(d_{iw}) \quad (9)$$

When we transform the constants and parameters, the equation is simplified into (10):

$$V_i = \mu^o \cdot \ln d_{oi} + \mu^w \cdot \ln d_{iw} + v \quad (10)$$

In which both μ^o and μ^w are assumed to be negative, because of the negative relation between distance and residential location choice. They are relation specific parameters, which can be estimated with a MNL function. The constant v can be eliminated when using the Logit method. The estimation is presented in the next chapter.

6 Estimation of the simultaneous influence

In this chapter we will estimate the simultaneous influence of migration and commuting distance with a Multinomial Logit Model. But first we will present the generation of a choice set generation based on the migration and commuting distance.

6.1 Initial choice set generation: residential search area

As mentioned in the description of the conceptual model, we will use systematic choice sets to estimate the choice model. For the first attempts, we only incorporate spatial restrictions to form the choice set, a so-called search area. We assume that, given the strong influence of the residential and work location, these spatial relations build up the search area.

The search area, from which randomly a certain number of locations can be chosen to construct the choice set, is supposed to include the locations around the residential and the work location as well as the locations between. After all, when looking for a new residential location, a household also takes into account the dwellings that have an even distance from the (primary) workplace and their former residential location. This implies that plain circles around both locations are not representing the actual situation well. This is why we introduce the ellipse as a geometric form to construct the search area.

An ellipse is built up by two fixed points (the foci), which are separated by a distance. From these foci two radii can be drawn, which's sum is a constant. In our case the foci are the (former) residential location and work location and the distance is the (former) commuting distance. The two radii represent the migration distance and the (new) commuting distance. The sum of the radii is a kind of upper boundary for the size of the ellipse. Figure 3 gives an image of the ellipse.

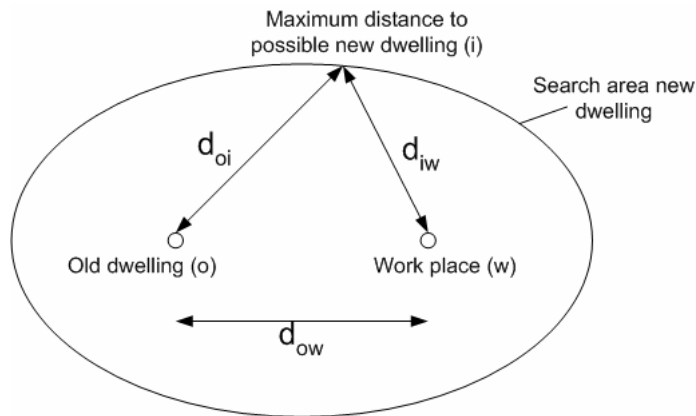


FIGURE 3: Geometric formulation Ellipse.

In order to determine which locations belong to the search area, we composed the following expression (8),

$$d_{oi} + d_{iw} - d_{ow} \leq E_{\max} \quad (8)$$

with: d_{oi} = distance between old dwelling o and optional, new dwelling i
 d_{iw} = distance between optional, new dwelling i and workplace w
 d_{ow} = distance between old dwelling o and workplace w
 E_{\max} = parameter representing the maximum size of the ellipse

To determine the E_{\max} we calculated the E-values from the moves that were in the survey. When we look at the distribution of the E, we see that almost 97% of all moves have an E smaller than 60 minutes. When we assume that the remaining 3% can be considered as extreme outliers, 60 minutes is a good value for the E_{\max} .

With the ellipse approach, we have a generic method to determine whether a location is part of the search area of a household with a given work and residential location. To illustrate the ellipse concept we have depicted some expressive examples in figure 4. From this picture we can see that not only the vicinities of the current residential or work location belong to the optional locations, but the zones in between as well. Furthermore the effect of the use of distance in term of (peak) travel times can be noticed. Especially when we look at the zones more or less isolated from the other areas. This is probably due to the presence of a highway connection nearby.

This ellipse method is only a first attempt to construct case specific choice sets. It incorporates a spatial dimension in the choice set formation. However, some remarks can be made. First of all, other, non-spatial issues play a part when developing a representative choice set for a residential choice problem. After all, the characteristics of the house, like its price, size and neighbourhood are decisive. Second, the proposed methodology doesn't account for a situation two-earners. That is, in the presented case the applied work location is that of the main wage earner of the household. Further in depth analysis of how the role of the other earners can be taken into account will be studied in next research steps.

The actual choice set that we have used to estimate the discrete choice model is drawn from the entire supply of dwellings per 4-digit postal zone. This means that a zone with more vacant dwellings has a larger chance to be chosen.

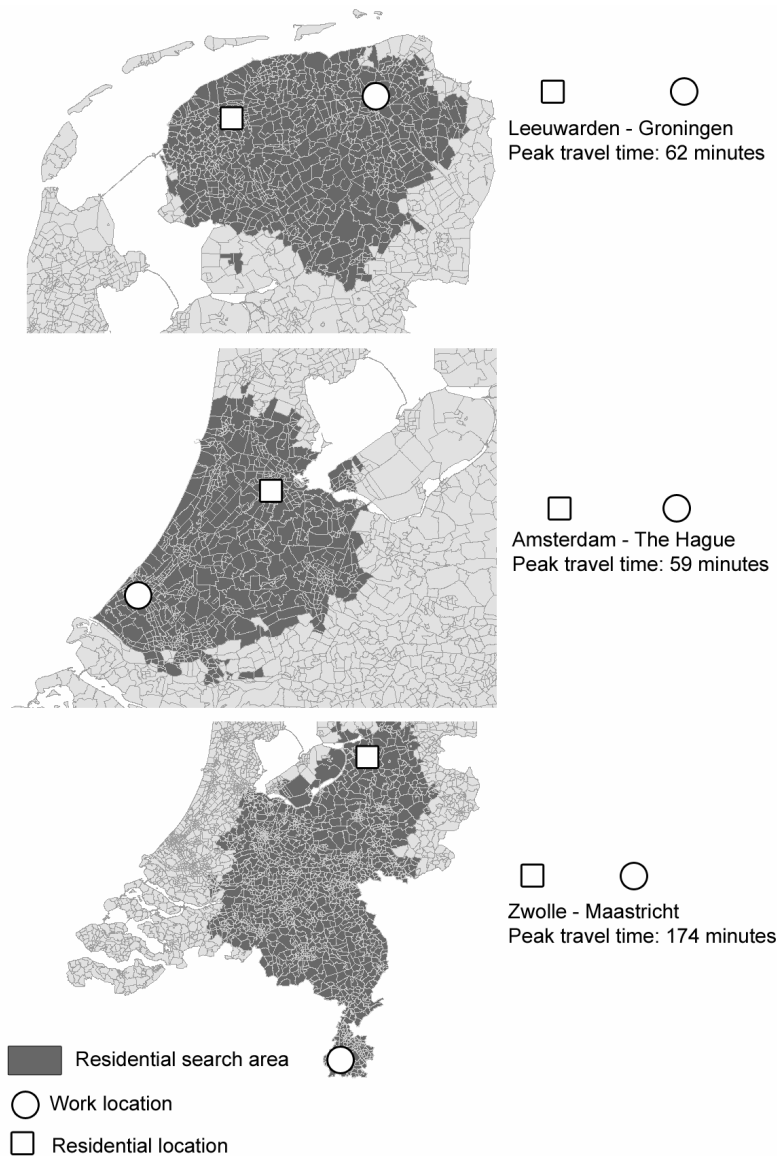


FIGURE 4: Examples search area.

6.2 Estimation of the MNL

In chapter five we have derived a utility function for residential location choice. In this function only the commuting and migration distances are incorporated, according to the negative power distance decay function. With the method presented in the previous section, we have constructed a choice set in which, next to the chosen alternative nine other alternatives are included. These alternatives are actual dwellings, from the supply present at the moment of choice. We have used the Biogeme software package to estimate a MNL model (Bierlaire et al., 2003).

The results are shown in table 4. From these results it can be seen that the simultaneous model has quite a large explanatory power (Rho-square = 0.428) and that both distances are significant and that the sign of the parameters is negative as expected. With a

parameter value of -1.910 the migration distance (μ^o) is almost twice as influential as the commuting distance ($\mu^w = -1.074$). This means that a 1% increase in commuting distance results in almost 2% decrease of chance that the alternative is chosen. For commuting distance the same increase reduces the chance by 1%.

TABLE 4: Parameter values estimated MNL model with negative power distance decay functions for migration and commuting distance.

Parameter	Value	t-test
μ^o	-1.91	-8.65e+01
μ^w	-1.07	-4.46e+01

Rho-square: 0.428

6.3 Conclusion estimation MNL

As can be seen from the estimation results, the negative power function gives the best approximation. With a relatively high Rho-square we have a confident base for further research. However, regarding the theoretical choice behaviour, it is well possible that combinations with other power distance decay functions give better results. After all, when the former dwelling and work place are fairly far apart from each other, it is likely that a location in between these locations is found most suitable. This behaviour is better represented by a function with a Log-logistic form than a power, because the latter one has a very steep decrease in likelihood, reducing the chances of options with a larger migration distance strongly.

7 Resume & Conclusion

We have presented an approach to model the residential choice behaviour of households in integrated land-use and transportation models. Within this approach special attention is given to the influence of accessibility on this behaviour. From the literature it appears that many (empirical) researches were not quite able to measure this influence. This is probably due to the fact that accessibility, when incorporated in the analysis, is overshadowed in importance by many other residential choice factors. New insights however, state that more individual accessibility measures (like the personal commuting time and migration distance) seem to be more significant and measurable variables.

The conceptual model presented in this paper consists of three main steps: the use of specific groups of households, which have a certain, common preference for the type of dwelling, living environment and mobility services. The second step aims on the application of systematic choice sets. After it is determined to what destination group a household belongs to, a large amount of alternatives can be left out because they are not acceptable for this type of household. Finally, the last step is a disaggregated discrete choice model, in which the personal attributes of a household are incorporated.

As for the ways that accessibility can be built into the model, we tend to make a distinction between personal and generic accessibility. The latter often turns out to have a weak influence on residential choice. On the other hand, individual measures like migration distance, commuting distance and the proximity of household type-specific services, matter to the household itself when it makes a decision where to settle.

In the last part of the paper we presented the first attempts to apply the concept of personal accessibility. We derived a utility function wherein the migration distance and commuting time are integrated. When regarded separately, for both motives the distance decay relations are estimated best by negative power functions. However, from the migration figures it can be seen that these distances work simultaneous in influencing the residential location choice. That is why we estimated Multinomial Logit Models in chapter 6 to determine the relative importance of commuting and migration distance. The choice set used for MNL-model is a random sample from the supply of dwelling that are located within an ellipse formed search area around the former residential location and the work place.

The results of these estimations show that when using a negative power, the influence of migration distance on residential location choice is almost twice as strong as the commuting distance. Also it turns out that an increase in migration distance of 1% reduces the chance of an alternative to be chosen with almost 2%.

7.1 Next research steps

After the first attempts of estimating the simultaneous influence of migration and commuting distance on residential choice behaviour, we want to experiment more with combinations of different distance decay functions. Like mentioned in chapter six, it is likely that other formulation than the power function are more suitable from a theoretical point of view.

In next research steps the Logit model can be extended with more variables regarding the characteristics of the dwelling, location/neighbourhood and household. We expect that in particular the dwelling attributes like size, tenure, type and price will add a considerable amount of explaining potential.

Also possibilities to experiment with the formulation of the discrete choice model are regarded. For example, it could be beneficial to try a nested Logit approach, in order to nest the characteristics of the location and the dwelling. Other improvements in the near future include constructing a search area for two-earners and using groups of households. Finally, we strive for a model as discussed in the previous section on the conceptual choice model.

Acknowledgements

The presented research is made possible by the support of other parties. We would like to thank the following institutes for their contribution in means of knowledge, data and software:

- The Netherlands Institute for Spatial research (Ruimtelijk Planbureau), Den Haag, Netherlands.
- ABF Research, Delft, The Netherlands
- Geo-Database Management Centre from the department Geodetic Engineering of Delft University of Technology, Delft, The Netherlands.

References

- Alonso, W. (1964) Location and land use: toward a general theory of land rent, Harvard University Press, Cambridge, Massachusetts, USA.
- Axhausen, D.M. Scott, A. König and Jürgens (2001) Locations, commitments and activity spaces, Presented at the Survive Workshop, Bonn, December 2001, Arbeitbericht Verkehrs- und Raumplanung **96**, IVT, ETH Zürich.

- Axhausen, K.W. & A. König, *Mobilitäts Werkzeuge und Wohnstandorte: Mobiplan stated choice – Experimente*, Vortrag bei der AMUS 2001 Aachen, Arbeitsbericht Verkehrs- und Raumplanung **79**, ETH, Zürich, Switzerland.
- Benjamin, J. and R.E. Paaswell (1981) A psychometric analysis of residential location, *Socio-Economic Planning Sciences* **15** (6): 305-319.
- Bierlaire, M., Bolduc, D. and Godbout, M.-H. (2003) An introduction to BIOGEME (Version 0.7), URL: roso.epfl.ch/mbi/biogeme/doc/tutorial.pdf
- Blijie, H.P. (2004) 'A New Theoretical Model for Residential Choice Behaviour and Measuring the Influence of Accessibility', presented at the 17th European Advanced Studies Institute in Regional Science, June 28th- July 4th 2004, Split, Croatia.
- Brun, J. & J. Fagnini (1994) Lifestyles and Locational Choices – Trade-offs and Compromises: A Case-study of Middle-class Couples Living in the Ile-de-France Region, *Urban Studies* **31**, no.6, pp 921-934.
- Christaller, W. (1962) 'Die Hierarchie der Städte', Proceedings of the IGU Symposium in Urban Geography, Lund, 1960. (Knut Norborg, ed.), Lund Studies in Geography, Series B, Human Geography, **24**, pp. 3-11.
- Clark, W.A.V. & F.M. Dieleman (1996) *Households and housing: choice and outcomes in the housing market*, New Brunswick, New Jersey, New York, USA.
- Clark, W.A.V. and W.F.J. Van Lierop (1986), 'Residential mobility and household location modelling', *Handbook of regional and urban economics: Volume I* (ed. P. Nijkamp), Elsevier Science Publishers.
- Evers, G.H.M. (1990) The residential location and workplace choice: a nested multinomial logit model. In: *Spatial Choices and Processes*, editors: M.M. Fischer, P. Nijkamp and Y.Y. Papageorgiou, Elsevier Science Publishers, B.V. (North-Holland), 313-329.
- Fotheringham, A. S. and M. E. O'Kelly (1989). *Spatial Interaction Models: Formulations and Applications*. Dordrecht, Kluwer Academic Publishers.
- Geurs, K.T. and Ritsema van Eck J.R. (2001) *Accessibility measures: review and applications*, RIVM Report 408505006, National Institute of Public Health and Environment, Bilthoven, The Netherlands.
- Hensher, D.A. & K.J. Button (2000) *Handbook of transport modelling*, Pergamon, Amsterdam, Netherlands.
- Hooimeijer, P. (1990) 'Towards a special demography of housing'. In: C. Hazeau en G. Frinking (eds.) *Emerging issues in demographic research*, Elsevier Science Publishers, p. 281-295, Amsterdam, Netherlands.
- Krizek, K.J. & P. Waddell (2002) 'Analysis of Lifestyle Choices: Neighborhood Type, Travel Patterns, and Activity Participation', *Transportation research record* **1807**, pp. 119-128.
- Lowry, I.S. (1964) *A model of Metropolis*, RM-4035-RC, Rand Corporation, Santa Monica, California, USA.
- McFadden, D. (1978). Modelling the choice of residential location, in A. Karlquist et al. (ed.), *Spatial interaction theory and residential location*, North-Holland, Amsterdam, pp. 75-96.
- Miller, E.J., D.S. Kriger and J.D. Hunt (1999), 'A Research and Development Program for Integrated Urban Models', *Transportation Research Record* **1685**, 1999, pp. 169-176.
- Molin, E.J.E., and H.J.P Timmermans (2003), *Transport considerations in residential choice decisions: accumulated evidence from the Benelux*, in: *Proceedings of the 82-th Annual Meeting of the Transport Research Board*, Washington, D.C..
- Muth, R.F. (1969) *Cities and housing. The spatial pattern of urban residential land use*, University of Chicago, Chicago, Illinois, USA.
- Nijkamp, P. and A. Reggiani (1992) *Interaction, evolution and chaos in space*, Springer, Berlin, Germany.

- Olatubara, C.O. (1998) 'An Alternative Approach to the Urban Residential Location Decision in Nigeria: The Nestling Idea', *Habitat international* **22**, pp. 57-67.
- Pagliari, F. and J. Preston (2003) *The Impact of transport on residential location - Final report*, Transport Studies Unit, University of Oxford, Oxford, UK.
- Pinkster, F & R. van Kempen (2002) *Leefstijlen en woonmilieuvorkeuren*, Urban and Regional research centre Utrecht, Faculteit Ruimtelijke Wetenschappen, Universiteit Utrecht, Utrecht, The Netherlands.
- Priemus, H. (1984) *Verhuistheorieën en de verdeling van de woningvoorraad*, Delftse Universitaire Pers, Delft, Netherlands.
- Rich, J.H. and Nielsen, O.A. (2001): 'A microeconomic model for car ownership, resident and work location'. European Transport Conference (PTRC). CDROM with proceedings, PTRC, Cambridge, September.
- Robson, B.T. (1975) *Urban social areas, ser.: Theory and Practice in Geography*, Oxford University Press, London, UK.
- Rosen, S. (1974) 'Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition'. *Journal of Political Economy* **82**: 34-55.
- Rossi, P.H. (1955), *Why Families move; a study in the social psychology of urban residential mobility*, The Free Press, Glencoe, Illinois, USA.
- Rouwendaal, J. (1989) *Choice and allocation models for the housing market, ser.: Studies in operational region sciences*, Kluwer Academic Publishers, Dordrecht, Netherlands.
- Srour, I.A., K.M. Kockelman and T.P. Dunn (2002) 'Accessibility Indices: A Connection to Residential Land Prices and Location Choices', Presented at the 81st Annual Meeting of the Transportation Research Board, Washington, USA.
- Timmermans, H., A. Borgers, J. van Dijk and H. Opperwal (1992) 'Residential choice behaviour of dual earner households: a decompositional joint choice model', *Environment and Planning A*, **24** (3): 517-533.
- Van der Waerden, P, A. Borgers and H. Timmermans (2004) 'Choice Set Composition in the Context of Pedestrians' Route Choice Modeling', presented at the 83rd Annual Meeting of the Transportation Research Board, Washington, D.C., USA.
- Waddel, P. (1996), *Accessibility and Residential Location: The Interaction of Workplace, Residential Mobility, Tenure, and Location Choices*, presented at the 1996 Lincoln Land Institute TRED Conference, Cambridge, Massachusetts.
- Walmsley, D. J. and Lewis, G.J. (1993) *People and Environment: Behavioural Approaches in Human Geography*, 2nd edition, John Wiley and Sons, New York, New York, USA.
- Weber, A. (1909) *Über den Standort der Industrien, Teil 1: Reine Theorie des Standorts*, Tübingen, Mohr, Germany.
- Wegener, M. and F. Fürst (1999), *Land-Use Transport Interaction: State of the Art*, TRANSLAND Integration of Transport and Land Use Planning Deliverable D2a, University of Dortmund, Dortmund, Germany.
- Weisbrod, G.E., Lerman, S.R., & Ben-Akiva, M. (1980), Tradeoffs in residential location decisions: transportation versus other factors. *Transport Policy and Decision Making*, 1, 13-26.
- Young, W. (1984) 'Modelling Residential Location Choice', *Australian Geographer* **16**, pp. 21-28.